

# Detecção de Objetos Através de Rede Neural do tipo FOMO

---

## OBJETIVO

O presente projeto tem como objetivo explorar o pipeline de desenvolvimento das seguintes técnicas:

- Implementação de Algoritmo de Detecção de Objetos em microcontrolador através de modelo de rede neural compacta do tipo **FOMO<sub>1</sub>** (utilizando-se os seguintes frameworks de ML: **Tensorflow, Tensorflow Lite e Tensorflow Lite Micro**).
- Implementação e testes de interface de câmeras tipo webcam em microcontrolador por intermédio de porta USB em modo host (utilizando-se stack **UVC<sub>2</sub> : USB Video Device Class**).
- Implementação e testes de algoritmo decodificador **JPEG** em microcontrolador (visando a decodificação de frames de imagens **MJPEG** das webcams para formato raw/matricial, próprio para processamento numérico pelos frameworks de ML).

## METODOLOGIA

O projeto será implementado sobre a plataforma de hardware **ESP32-S3-EYE** (SoC ESP32-S3, 8 MB de memória Flash, 8 MB de memória PSRAM) e câmera externa (webcam). Utilizar-se-á ambas as câmeras (CMOS interna e USB externa) para captura de imagens, sendo a comparação (de aplicabilidade e desempenho) delas um dos objetivos secundários do projeto.

O estudo em torno da implementação de modelo de rede neural compacta tipo **FOMO<sub>1</sub> (Faster Objects, More Objects)** será o objetivo principal. Este modelo de rede neural foi inicialmente proposto pela **Edge Impulse<sub>3</sub>** e por **Mat Palm<sub>4</sub>**.

A maioria dos modelos de rede neural para detecção de objetos possuem requisitos computacionais e de memória muito superiores às capacidades dos microcontroladores de uso geral.

Um modelo tipo **FOMO<sub>1</sub>**, por outro lado, requer apenas algumas centenas de kilobytes de memória, o que o torna

# Detecção de Objetos Através de Redes Neurais do tipo FOMO

---

uma ferramenta promissora para implementação de algoritmos detectores de objetos em microcontroladores (resource constrained devices).

A plataforma **Edge Impulse<sub>3</sub>** será utilizada como referência de design para acelerar o início de desenvolvimento do modelo de rede neural **FOMO<sub>1</sub>**.

A modelo inicial gerado pela plataforma (em código Tensorflow/Keras) será exportado para ser investigado em detalhes, externamente, em instância Google Colab ou desktop local com GPU.

Serão avaliadas os dois caminhos de implantação (deployment) modelo de rede neural na plataforma de hardware **ESP32-S3-EYE**:

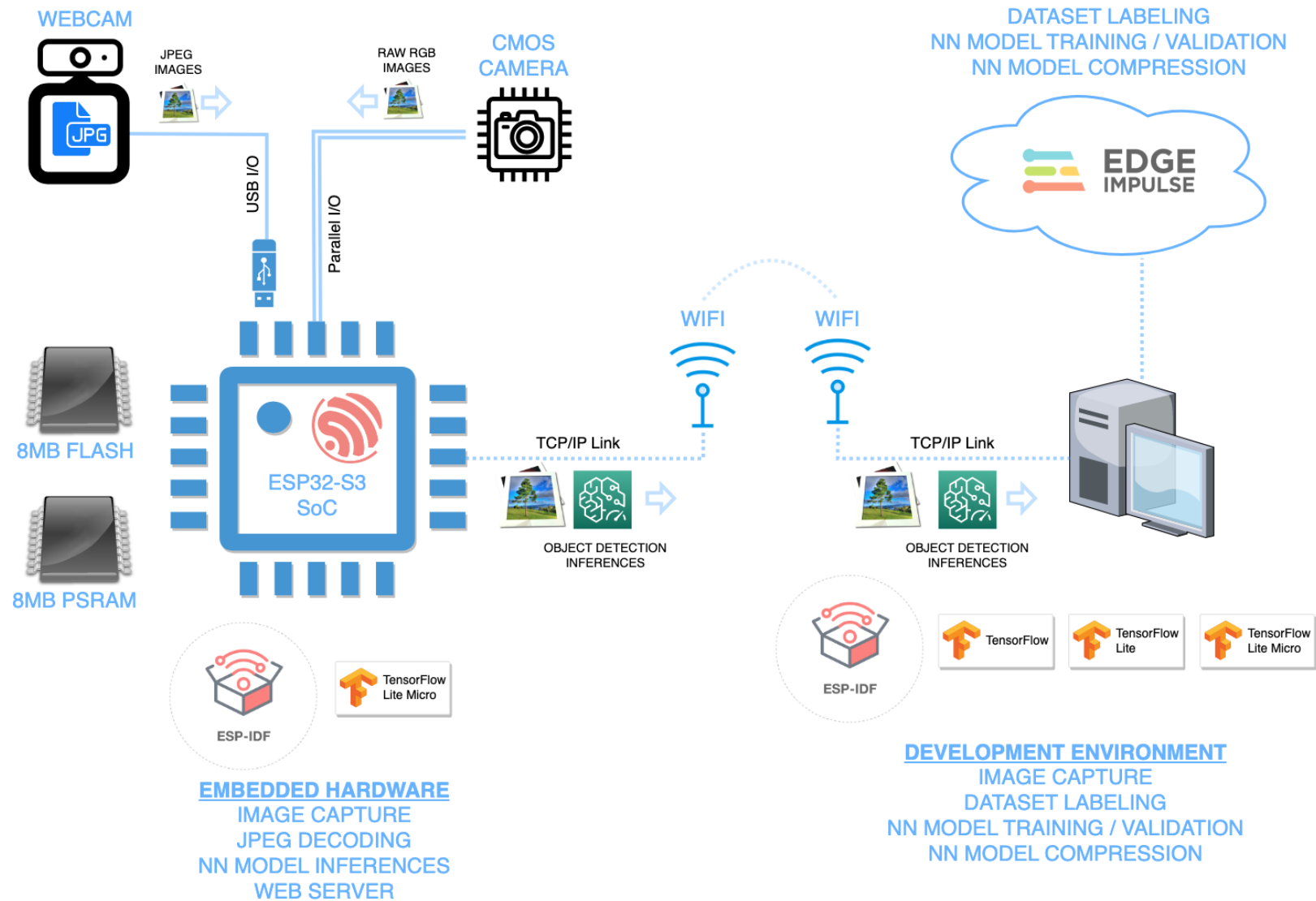
- Modelo **FOMO<sub>1</sub>** treinado via **Edge Impulse<sub>3</sub>**;
- Modelo alternativo baseado no **FOMO<sub>1</sub>**, mas treinado localmente via **TensorFlow**, visando um maior controle sobre o treinamento, inclusão de camadas regularização (tais como Dropout, Regularizações L1 e L2 ), aplicação

de mais opções de “data augmentation” nos datasets de treinamento além de outros recursos de otimização da performance do modelo.

Classificação de imagem é legal. No entanto, detecção de objetos (localização e contagem) é o que realmente importa na maioria das aplicações práticas de visão computacional.

A idéia de se utilizar um microcontrolador para a captura de imagens de uma webcam (ou até mesmo várias câmeras), além de processar a detecção de objetos localmente é a motivação e o desafio deste projeto. A detecção de objetos através de processamento distribuído (em microcontrolador de baixo custo) tem um grande potencial de aplicação na indústria.

# Detecção de Objetos Através de Redes Neurais do tipo FOMO



# Detecção de Objetos Através de Redes Neurais do tipo FOMO

---

## REFERÊNCIAS

[1] **“Constrained Object Detection on Microcontrollers with FOMO”**, Edge Impulse (Shawn Hymel)

▶ [https://cms.tinymml.org/wp-content/uploads/summit2022/tinyML\\_Talks\\_Shawn\\_Hymel\\_220405.pdf](https://cms.tinymml.org/wp-content/uploads/summit2022/tinyML_Talks_Shawn_Hymel_220405.pdf)

▶ <https://edgeimpulse.com/fomo>

[2] **UVC : USB Video Device Class**

▶ <https://www.usb.org/document-library/video-class-v11-document-set>

[3] **Edge Impulse** : O Edge Impulse é uma plataforma de desenvolvimento de aprendizado de máquina voltada para dispositivos de borda (edge devices).

▶ <https://edgeimpulse.com/>

[4] **Counting Bees on a Rasp Pi with a Conv Net (Mat Palm)**

▶ [https://matpalm.com/blog/counting\\_bees/](https://matpalm.com/blog/counting_bees/)

[5] **TechTalks : FOMO is a TinyML neural network for real-time object detection**

▶ <https://bdtechtalks.com/2022/04/18/fomo-tinymml-object-detection/>

[6] **“FOMO: Real-Time Object Detection on Microcontrollers,” a Presentation from Edge Impulse (Jan Jongboom)**

▶ <https://www.edge-ai-vision.com/2022/06/fomo-real-time-object-detection-on-microcontrollers-a-presentation-from-edge-impulse/>

▶ [https://www.edge-ai-vision.com/wp-content/uploads/2022/05/E3W04\\_Jongboom\\_Edge\\_Impulse\\_2022.pdf](https://www.edge-ai-vision.com/wp-content/uploads/2022/05/E3W04_Jongboom_Edge_Impulse_2022.pdf)